

2001-65637

提出日 平成13年 3月 8日
頁: 1/ 2

整理番号=2033830017

【書類名】 特許願

【整理番号】 2033830017

【提出日】 平成13年3月8日

【あて先】 特許庁長官殿

【国際特許分類】 G10L 3/00

【発明者】

【住所又は居所】 大阪府門真市大字門真1006番地 松下電器産業株式
式会社内

【氏名】 脇田 由実

【発明者】

【住所又は居所】 大阪府門真市大字門真1006番地 松下電器産業株式
式会社内

【氏名】 松井 謙二

【特許出願人】

【識別番号】 000005821

【氏名又は名称】 松下電器産業株式会社

【代理人】

【識別番号】 100092794

【弁理士】

【氏名又は名称】 松田 正道

【電話番号】 06-6397-2840

【先の出願に基づく優先権主張】

【出願番号】 特願2000-66494

【出願日】 平成12年3月10日

【手数料の表示】

【予納台帳番号】 009896

【納付金額】 21000

【提出物件の目録】

【物件名】 明細書 1

整理番号＝ 2 0 3 3 8 3 0 0 1 7

提出日 平成 1 3 年 3 月 8 日
頁: 2 / 2

【物件名】	図面	1
【物件名】	要約書	1
【包括委任状番号】		9 0 0 6 0 2 7
【プルーフの要否】	要	

【書類名】 明細書

【発明の名称】 表現変換方法及び表現変換装置

【特許請求の範囲】

【請求項1】 コーパスの各文について、その文から重要語を選んで、それらの重要語のうちの、所定の数の組み合わせの中から共起関係のある重要語の組み合わせを特定して、その特定した重要語の組み合わせとその重要語を選択する対象となった文の変換後の表現とを予め対応させておき、

入力されてきた文から予め決まっている重要語を選んで、組み合わせ、その組み合わせた重要語の組み合わせと、前記各文の予め前記特定した重要語の組み合わせとを対比し、その対比結果が一致もしくは類似度の高い文を1または2以上選んで、その文の変換後の表現を出力する表現変換方法。

【請求項2】 予め重要語を所定の性質に基づきクラス化してクラス名をつけておいたクラス化情報を利用して、コーパスの各文について、その文から重要語を選んで、それらの各重要語が属するクラスのうちの、所定の数の組み合わせの中から共起関係のあるクラスの組み合わせを特定して、その特定したクラスの組み合わせと前記重要語を選択する対象となった文の変換後の表現とを予め対応させておき、

入力されてきた文から予め決まっている重要語を選んで、それらの各重要語が属するクラスを組み合わせ、その組み合わせたクラスの組み合わせと、前記各文の予め前記特定したクラスの組み合わせとを対比し、その対比結果が一致もしくは類似度の高い文を1または2以上選んで、その文の変換後の表現を出力する表現変換方法。

【請求項3】 コーパスの各文について、その文から重要語を選んで、それらの重要語のうちの、所定の数の組み合わせの中から共起関係のある重要語の組み合わせを特定して、その特定した重要語の組み合わせとその重要語を選択する対象となった文の変換後の表現とを予め対応させておき、

予め重要語を所定の性質に基づきクラス化してクラス名をつけておいたクラス化情報を利用して、前記特定した重要語の組み合わせをクラスの組み合わせに対応させることによって、前記文のクラスの組み合わせを特定しておき、

入力されてきた文から予め決まっている重要語を選んで、それらの各重要語が属するクラスを組み合わせ、その組み合わせたクラスの組み合わせと、前記各文の予め前記特定したクラスの組み合わせとを対比し、その対比結果が一致もしくは類似度の高い文を1または2以上選んで、その文の変換後の表現を出力する表現変換方法。

【請求項4】 コーパスの各文について、その文から重要語を選んで、それらの重要語のうちの、所定の数の組み合わせの中から共起関係のある重要語の組み合わせを特定して、その特定した重要語の組み合わせとその重要語を選択する対象となった文の変換後の表現とを予め対応させておく対応付け手段と、

入力されてきた文から予め決まっている重要語を選んで、組み合わせ、その組み合わせた重要語の組み合わせと、前記各文の予め前記特定した重要語の組み合わせとを対比し、その対比結果が一致もしくは類似度の高い文を1または2以上選んで、その文の変換後の表現を出力する変換手段とを備えた表現変換装置。

【請求項5】 前記変換手段は、前記対比結果が類似度の高い場合、不一致となった方の、重要語の組み合わせに属する重要語で、一致している方の、重要語の組み合わせに含まれていない重要語については、その変換後の部分を前記選ばれた表現から取り除いて前記表現を出力する請求項4記載の表現変換装置。

【請求項6】 前記変換後の表現は、重要語または重要語の対訳単語のみで構成されていることを特徴とする請求項4記載の表現変換装置。

【請求項7】 予め重要語を所定の性質に基づきクラス化してクラス名をつけておいたクラス化情報を利用して、コーパスの各文について、その文から重要語を選んで、それらの各重要語が属するクラスのうちの、所定の数の組み合わせの中から共起関係のあるクラスの組み合わせを特定して、その特定したクラスの組み合わせと前記重要語を選択する対象となった文の変換後の表現とを予め対応させておく対応付け手段と、

入力されてきた文から予め決まっている重要語を選んで、それら各重要語が属するクラスを組み合わせ、その組み合わせたクラスの組み合わせと、前記各文の予め前記特定したクラスの組み合わせとを対比し、その対比結果が一致もしくは類似度の高い文を1または2以上選んで、その文の変換後の表現を出力する変換

手段とを備えた表現変換装置。

【請求項８】 コーパスの各文について、その文から重要語を選んで、それらの重要語のうちの、所定の数の組み合わせの中から共起関係のある重要語の組み合わせを特定して、その特定した重要語の組み合わせとその重要語を選択する対象となった文の変換後の表現とを予め対応させておき、

予め重要語を所定の性質に基づきクラス化してクラス名をつけておいたクラス化情報を利用して、前記特定した重要語の組み合わせをクラスの組み合わせに対応させることによって、前記文のクラスの組み合わせを特定しておく対応付け手段と、

入力されてきた文から予め決まっている重要語を選んで、それら各重要語が属するクラスを組み合わせ、その組み合わせたクラスの組み合わせと、前記各文の予め前記特定したクラスの組み合わせとを対比し、その対比結果が一致もしくは類似度の高い文を１または２以上選んで、その文の変換後の表現を出力する変換手段とを備えた表現変換装置。

【請求項９】 前記変換手段は、前記対比結果が類似度の高い場合、不一致となった方の、クラスの組み合わせに属するクラスで、一致している方の、クラスの組み合わせに含まれていないクラスについては、その変換後の部分を前記選ばれた表現から取り除いて前記表現を出力する請求項７または８に記載の表現変換装置。

【請求項１０】 前記変換後の表現は、クラスのみから構成されている請求項７または８に記載の表現変換装置。

【請求項１１】 請求項４記載の表現変換装置の、コーパスの各文について、その文から重要語を選んで、それらの重要語のうちの、所定の数の組み合わせの中から共起関係のある重要語の組み合わせを特定して、その特定した重要語の組み合わせとその重要語を選択する対象となった文の変換後の表現とを予め対応させておく対応付け手段と、

入力されてきた文から予め決まっている重要語を選んで、組み合わせ、その組み合わせた重要語の組み合わせと、前記各文の予め前記特定した重要語の組み合わせとを対比し、その対比結果が一致もしくは類似度の高い文を１または２以上

選んで、その文の変換後の表現を出力する変換手段との全部または一部としてコンピュータを機能させるためのプログラム。

【請求項12】 請求項7記載の表現変換装置の、予め重要語を所定の性質に基づきクラス化してクラス名をつけておいたクラス化情報を利用して、コーパスの各文について、その文から重要語を選んで、それらの各重要語が属するクラスのうちの、所定の数の組み合わせの中から共起関係のあるクラスの組み合わせを特定して、その特定したクラスの組み合わせと前記重要語を選択する対象となった文の変換後の表現とを予め対応させておく対応付け手段と、

入力されてきた文から予め決まっている重要語を選んで、それらの各重要語が属するクラスを組み合わせ、その組み合わせたクラスの組み合わせと、前記各文の予め前記特定したクラスの組み合わせとを対比し、その対比結果が一致もしくは類似度の高い文を1または2以上選んで、その文の変換後の表現を出力する変換手段との全部または一部としてコンピュータを機能させるためのプログラム。

【請求項13】 請求項8記載の表現変換装置の、コーパスの各文について、その文から重要語を選んで、それらの重要語のうちの、所定の数の組み合わせの中から共起関係のある重要語の組み合わせを特定して、その特定した重要語の組み合わせとその重要語を選択する対象となった文の変換後の表現とを予め対応させておき、

予め重要語を所定の性質に基づきクラス化してクラス名をつけておいたクラス化情報を利用して、前記特定した重要語の組み合わせをクラスの組み合わせに対応させることによって、前記文のクラスの組み合わせを特定しておく対応付け手段と、

入力されてきた文から予め決まっている重要語を選んで、それらの各重要語が属するクラスを組み合わせ、その組み合わせたクラスの組み合わせと、前記各文の予め前記特定したクラスの組み合わせとを対比し、その対比結果が一致もしくは類似度の高い文を1または2以上選んで、その文の変換後の表現を出力する変換手段との全部または一部としてコンピュータを機能させるためのプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、入力されてくる文の表現を変換して出力する表現変換方法、表現変換装置、及びプログラムに関するものであり、例えば、翻訳または通訳などの言語変換、書き言葉を話し言葉に変換する文型変換、複雑なまたは冗長な文章を要約して出力する要約文作成などを行う表現変換方法、表現変換装置、及びプログラムに関するものである。

【０００２】**【従来の技術】**

以下、従来の技術を通訳ソフトを例にして説明する。

【０００３】

通訳ソフトは、音声認識手段と言語翻訳手段とから構成されており、音響信号として入力された発声文を単語テキスト列で表示された出力文に変換するための音声認識と、単語テキスト列で表示された文を入力し他言語文に翻訳する表現変換としての言語翻訳とを順次実行することで通訳を実現している。

【０００４】

さらに、上記のようにして言語翻訳を実行する言語翻訳手段は、入力文の統語的または意味的構造を解析する言語解析手段と、解析結果に基づいて他言語に変換する言語変換手段と、翻訳結果から自然な出力文を生成する出力文生成手段とから構成されている。

【０００５】

通訳ソフトや一部の翻訳ソフトでは、話し言葉に頻繁に見られるような、文法規則から外れたフランクな表現を正しく翻訳するために、文法規則に従った言語解析を行うのではなく、実際に入力される文と類似した用例文を学習しておき、類似用例を検索することで言語解析を行う手法を使用している場合が多い。以下、このような従来の通訳ソフト例を図８の例に沿って説明する。

【０００６】

ここでは、日本語の発声文例を英語の発声文例に通訳する場合について説明する。また、通訳される前の言語を原言語、通訳された後の言語を目的言語と呼ぶことにする。

【0007】

通訳を行う前に、図8の対訳コーパス1から発声文規則を抽出する。対訳コーパス1には、対訳となる日本語の発声文例と英語の発声文例との対である対訳発声文例が複数記述されている。図9の(a)に対訳コーパス1に記述されている対訳発声文例の例を対訳発声文例70として示す。

【0008】

ここでは、一部の単語が認識誤りを起したり省略して発声されることを考慮し、発声文例を意味的なまとまりとしての最小単位(以後、フレーズと呼ぶ)に分割し、フレーズ内規則とフレーズ間の依存規則とを作成している。

【0009】

まずフレーズ決手段61は、対訳発声文例をフレーズ単位に分割する。図9の(b)にこのようにして分割された対訳フレーズを対訳フレーズ(A)71、対訳フレーズ(B)72として示す。

【0010】

次に、対訳フレーズ辞書作成部62は、対応フレーズ辞書62をフレーズ内の内容語を変数化された形で作成する。

【0011】

例えば、図9の(a)に示す対訳発声文例70では、「部屋の予約をお願いしたいんですが」「I'd like to reserve a room」という発声文例が記述されており、これらをフレーズ単位に分割すると、対訳フレーズ(A)71として(A)「部屋の予約」「reserve a room」、対訳フレーズ(B)として(B)「お願いしたいんですが」「I'd like to」という(A)(B)2つの対訳フレーズに分割される。

【0012】

そして「部屋」「予約」「お願い」などの内容語は、図9の(e)に示すように予め作成されている分類語意表64を利用して各々XYZという変数で表される。ここで、分類語意表64とは、各変数の値として取りうる内容語が記述された一覧表である。例えば変数Xは、「部屋」、「会議室」、「車」などの値をとり、「部屋」という内容語は、変数Xが取りうる値になっている。従って、対訳フレ

ーズ (A) 71の「部屋」という内容語は変数Xに置き換えられる。

【0013】

このようにして、(A)「XのY」「YX」、(B)「Zしたいんですが」「I'd like to」という2つの対訳フレーズ規則が対訳フレーズ辞書62に記述される。

【0014】

また、各々のフレーズの順序関係を規則化するため、図9の(d)にフレーズ間規則63として示すように、「(A)を(B)」「(B)(A)」というフレーズ間の関係が、フレーズ間規則テーブル63に各々保管される。このような処理が対訳コーパス1に含まれた全発声文分について行われる。

【0015】

通訳の際には、まず発声された原言語音声は音声認識手段64に入力される。音声認識手段64では、たとえば、対訳フレーズ辞書62にフレーズとして記述されている単語列とフレーズ間規則63に記述されているフレーズ列とから予測される可能性のある単語列の中から、音響的に最も類似している単語列を音声認識候補として出力する。

【0016】

言語翻訳手段65は、このようにして認識された連続単語列を入力し、入力された連続単語列を対訳フレーズ辞書62に記載されているフレーズ列に変換し、各フレーズ列に相当するフレーズ間規則63を探索する。そして、各フレーズの対訳である目的言語フレーズと目的言語のフレーズ間規則とから、入力原言語認識結果文を目的言語文に変換する。

【0017】

変換された目的言語文は出力文生成手段66に入力され、文法的な不自然さを修正する。たとえば、代名詞や動詞や助動詞における3人称化や複数化や過去形化などの最適化や、文全体の構成の最適化などが行われる。修正後の目的言語翻訳結果文はたとえばテキストとして出力される。

【0018】

【発明が解決しようとする課題】

しかしながら、従来の技術の通訳ソフトでは、文法規則から外れた入力文に対応できる良さがある反面、多数の異なる対訳フレーズとその組み合わせが、そのまま規則として記述されるため、変換規則が煩雑で膨大になってしまい処理時間もかかってしまうという課題を有している。

【００１９】

また、出力文生成部にて行われる文法的なチェックにも、複雑な規則が必要となり、特にゼロ代名詞の補完などは完全に正しく補完できる技術がなく、時々誤った補完を行ってしまうという課題を有している。

【００２０】

また、音声認識誤りなどにより、部分的に誤っている文が言語翻訳部に入力された場合、誤ったフレーズ内およびフレーズ間規則に基づき言語変換がなされるため、全く意図が伝わらない翻訳結果を出力するという課題を有している。

【００２１】

本発明は、上記課題を考慮し、コンパクトな構成であり、かつ高速に処理可能な表現変換方法、表現変換装置、及びプログラムを提供することを目的とするものである。

【００２２】

また、本発明は、上記課題を考慮し、音声認識の誤りなどにより入力文中の重要語以外で認識誤りを起こしても、出力文の質に悪影響を及ぼすことなく意図が正しく伝わる結果を出力できる表現変換方法、表現変換装置、及びプログラムを提供することを目的とするものである。

【００２３】

また、本発明は、上記課題を考慮し、音声認識の誤りなどにより入力文中で認識誤りを起こしても、従来のように全く文意が伝わらない結果を出力することを回避することが出来る表現変換方法、表現変換装置、及びプログラムを提供することを目的とするものである。

【００２４】

【課題を解決するための手段】

上述した課題を解決するために、第１の本発明（請求項１に対応）は、コーパ

スの各文について、その文から重要語を選んで、それらの重要語のうちの、所定の数の組み合わせの中から共起関係のある重要語の組み合わせを特定して、その特定した重要語の組み合わせとその重要語を選択する対象となった文の変換後の表現とを予め対応させておき、

入力されてきた文から予め決まっている重要語を選んで、組み合わせ、その組み合わせた重要語の組み合わせと、前記各文の予め前記特定した重要語の組み合わせとを対比し、その対比結果が一致もしくは類似度の高い文を１または２以上選んで、その文の変換後の表現を出力する表現変換方法である。

【００２５】

また、第２の本発明（請求項２に対応）は、予め重要語を所定の性質に基づきクラス化してクラス名をつけておいたクラス化情報を利用して、コーパスの各文について、その文から重要語を選んで、それらの各重要語が属するクラスのうちの、所定の数の組み合わせの中から共起関係のあるクラスの組み合わせを特定して、その特定したクラスの組み合わせと前記重要語を選択する対象となった文の変換後の表現とを予め対応させておき、

入力されてきた文から予め決まっている重要語を選んで、それらの各重要語が属するクラスを組み合わせ、その組み合わせたクラスの組み合わせと、前記各文の予め前記特定したクラスの組み合わせとを対比し、その対比結果が一致もしくは類似度の高い文を１または２以上選んで、その文の変換後の表現を出力する表現変換方法である。

【００２６】

また、第３の本発明（請求項３に対応）は、コーパスの各文について、その文から重要語を選んで、それらの重要語のうちの、所定の数の組み合わせの中から共起関係のある重要語の組み合わせを特定して、その特定した重要語の組み合わせとその重要語を選択する対象となった文の変換後の表現とを予め対応させておき、

予め重要語を所定の性質に基づきクラス化してクラス名をつけておいたクラス化情報を利用して、前記特定した重要語の組み合わせをクラスの組み合わせに対応させることによって、前記文のクラスの組み合わせを特定しておき、

入力されてきた文から予め決まっている重要語を選んで、それらの各重要語が属するクラスを組み合わせ、その組み合わせたクラスの組み合わせと、前記各文の予め前記特定したクラスの組み合わせとを対比し、その対比結果が一致もしくは類似度の高い文を１または２以上選んで、その文の変換後の表現を出力する表現変換方法である。

【００２７】

また、第４の本発明（請求項４に対応）は、コーパスの各文について、その文から重要語を選んで、それらの重要語のうちの、所定の数の組み合わせの中から共起関係のある重要語の組み合わせを特定して、その特定した重要語の組み合わせとその重要語を選択する対象となった文の変換後の表現とを予め対応させておく対応付け手段と、

入力されてきた文から予め決まっている重要語を選んで、組み合わせ、その組み合わせた重要語の組み合わせと、前記各文の予め前記特定した重要語の組み合わせとを対比し、その対比結果が一致もしくは類似度の高い文を１または２以上選んで、その文の変換後の表現を出力する変換手段とを備えた表現変換装置である。

【００２８】

また、第５の本発明（請求項５に対応）は、前記変換手段は、前記対比結果が類似度の高い場合、不一致となった方の、重要語の組み合わせに属する重要語で、一致している方の、重要語の組み合わせに含まれていない重要語については、その変換後の部分を前記選ばれた表現から取り除いて前記表現を出力する第４の本発明に記載の表現変換装置である。

【００２９】

また、第６の本発明（請求項６に対応）は、前記変換後の表現は、重要語または重要語の対訳単語のみで構成されていることを特徴とする第４の本発明に記載の表現変換装置である。

【００３０】

また、第７の本発明（請求項７に対応）は、予め重要語を所定の性質に基づきクラス化してクラス名をつけておいたクラス化情報を利用して、コーパスの各文

について、その文から重要語を選んで、それらの各重要語が属するクラスのうちの、所定の数の組み合わせの中から共起関係のあるクラスの組み合わせを特定して、その特定したクラスの組み合わせと前記重要語を選択する対象となった文の変換後の表現とを予め対応させておく対応付け手段と、

入力されてきた文から予め決まっている重要語を選んで、それら各重要語が属するクラスを組み合わせ、その組み合わせたクラスの組み合わせと、前記各文の予め前記特定したクラスの組み合わせとを対比し、その対比結果が一致もしくは類似度の高い文を１または２以上選んで、その文の変換後の表現を出力する変換手段とを備えた表現変換装置である。

【００３１】

また、第８の本発明（請求項８に対応）は、コーパスの各文について、その文から重要語を選んで、それらの重要語のうちの、所定の数の組み合わせの中から共起関係のある重要語の組み合わせを特定して、その特定した重要語の組み合わせとその重要語を選択する対象となった文の変換後の表現とを予め対応させておき、

予め重要語を所定の性質に基づきクラス化してクラス名をつけておいたクラス化情報を利用して、前記特定した重要語の組み合わせをクラスの組み合わせに対応させることによって、前記文のクラスの組み合わせを特定しておく対応付け手段と、

入力されてきた文から予め決まっている重要語を選んで、それら各重要語が属するクラスを組み合わせ、その組み合わせたクラスの組み合わせと、前記各文の予め前記特定したクラスの組み合わせとを対比し、その対比結果が一致もしくは類似度の高い文を１または２以上選んで、その文の変換後の表現を出力する変換手段とを備えた表現変換装置である。

【００３２】

また、第９の本発明（請求項９に対応）は、前記変換手段は、前記対比結果が類似度の高い場合、不一致となった方の、クラスの組み合わせに属するクラスで、一致している方の、クラスの組み合わせに含まれていないクラスについては、その変換後の部分を前記選ばれた表現から取り除いて前記表現を出力する第７ま

たは8の本発明に記載の表現変換装置である。

【0033】

また、第10の本発明（請求項10に対応）は、前記変換後の表現は、クラス用語のみから構成されている第7または8の本発明に記載の表現変換装置である。

【0034】

また、第11の本発明（請求項11に対応）は、第4の本発明に記載の表現変換装置の、コーパスの各文について、その文から重要語を選んで、それらの重要語のうちの、所定の数の組み合わせの中から共起関係のある重要語の組み合わせを特定して、その特定した重要語の組み合わせとその重要語を選択する対象となった文の変換後の表現とを予め対応させておく対応付け手段と、

入力されてきた文から予め決まっている重要語を選んで、組み合わせ、その組み合わせた重要語の組み合わせと、前記各文の予め前記特定した重要語の組み合わせとを対比し、その対比結果が一致もしくは類似度の高い文を1または2以上選んで、その文の変換後の表現を出力する変換手段との全部または一部としてコンピュータを機能させるためのプログラムである。

【0035】

また、第12の本発明（請求項12に対応）は、第7の本発明に記載の表現変換装置の、予め重要語を所定の性質に基づきクラス化してクラス名をつけておいたクラス化情報を利用して、コーパスの各文について、その文から重要語を選んで、それらの各重要語が属するクラスのうちの、所定の数の組み合わせの中から共起関係のあるクラスの組み合わせを特定して、その特定したクラスの組み合わせと前記重要語を選択する対象となった文の変換後の表現とを予め対応させておく対応付け手段と、

入力されてきた文から予め決まっている重要語を選んで、それらの各重要語が属するクラスを組み合わせ、その組み合わせたクラスの組み合わせと、前記各文の予め前記特定したクラスの組み合わせとを対比し、その対比結果が一致もしくは類似度の高い文を1または2以上選んで、その文の変換後の表現を出力する変換手段との全部または一部としてコンピュータを機能させるためのプログラムで

ある。

【 0 0 3 6 】

また、第 1 3 の本発明（請求項 1 3 に対応）は、第 8 の本発明に記載の表現変換装置の、コーパスの各文について、その文から重要語を選んで、それらの重要語のうちの、所定の数の組み合わせの中から共起関係のある重要語の組み合わせを特定して、その特定した重要語の組み合わせとその重要語を選択する対象となった文の変換後の表現とを予め対応させておき、

予め重要語を所定の性質に基づきクラス化してクラス名をつけておいたクラス化情報を利用して、前記特定した重要語の組み合わせをクラスの組み合わせに対応させることによって、前記文のクラスの組み合わせを特定しておく対応付け手段と、

入力されてきた文から予め決まっている重要語を選んで、それらの各重要語が属するクラスを組み合わせ、その組み合わせたクラスの組み合わせと、前記各文の予め前記特定したクラスの組み合わせとを対比し、その対比結果が一致もしくは類似度の高い文を 1 または 2 以上選んで、その文の変換後の表現を出力する変換手段との全部または一部としてコンピュータを機能させるためのプログラムである。

【 0 0 3 7 】

次に本発明の動作を説明する。

【 0 0 3 8 】

本発明によれば、入力文から重要語を抽出し、抽出された重要語を用いて同じ文意を表す標準的または簡易的表現文に変換し出力することで、変換規則や文生成規則をコンパクトに構成することが可能となり簡易な処理で表現変換を行うことを可能にする。また、重要語以外の部分で誤った文が入力された場合でも、正しく文意が理解できる表現に変換でき、従来のように誤ったまま表現変換結果を出力するという問題を解決することが可能となる。

【 0 0 3 9 】

また、本発明によれば、入力文に含まれる一部の内容語の単語もしくはその変換された単語を重要語として抽出し、この重要語と入力文から推定される文意に

より決まる表現との組み合わせからなる標準的または簡易的表現文を生成することで、変換規則や文生成規則をコンパクトに構成することが可能となり簡易な処理で表現変換を行うことを可能にする。また、重要語以外の部分で誤った文が入力された場合でも、正しく文意が理解できる表現に変換でき、従来のように誤ったまま表現変換結果を出力するという問題を解決することが可能となる。

【００４０】

また、本発明によれば、入力文から重要語を抽出し、抽出された重要語の共起関係または共起関係から文意を推定し、重要語またはその対訳単語と推定された意図から予め決められている単語のみの組み合わせから標準的または簡易的表現を生成することで、変換規則や文生成規則をコンパクトに構成することが可能となり簡易な処理で表現変換を行うことを可能にする。また、重要語以外の部分で誤った文が入力された場合でも、正しく文意が理解できる表現に変換でき、従来のように誤ったまま表現変換結果を出力するという問題を解決することが可能となる。

【００４１】

また、本発明によれば、予め、出力される標準的または簡易的表現パターンとそのパターンに当てはまる重要語群及び重要語群間の共起関係または共起関係とを対応させている用例文データベースを作成しておき、入力文から重要語群を抽出し、抽出された重要語群と最も類似している重要語群をもつ表現パターンを用例文データベースから選択し、選択された用例文を出力することにより、上記本発明の効果に加え、現実に入力される文の種類やドメインや文型に忠実に確度よく表現変換を行うことを可能にする。

【００４２】

また、本発明によれば、用例文データベースに記載されている表現パターンが重要語または重要語の対訳単語のみで構成されていることにより、上記本発明の効果をさらに強化することが可能である。

【００４３】

また、本発明によれば、用例文データベースに記載されている表現パターン、重要語群及び重要語群間の共起関係または共起関係の重要語の記述を重要語を含む

単語クラスの記述とすることで、上記本発明の効果に加え、用例文データベースに含まれていない重要語が入力されても妥当な用例文を選択可能となり、より多種多様な入力文に対応できる表現変換が可能となる。

【0044】

また、本発明によれば、入力文から重要語群を抽出し、抽出された重要語間の関係から入力誤りである単語を推定し、推定された単語を除いた残りの重要語から文意を推定し、推定される文意により決まる単語の組み合わせから標準的または簡易的表現を生成することで、上記本発明に記載した効果に加えて、たとえ重要語が誤っていても、その誤りの重度に応じて、正しく文意が理解できる表現に変換するか、意味が理解できなかったことをユーザに告知することが可能となり、従来のように誤ったまま表現変換結果を出力するという問題を解決することが可能となる。

【0045】

【発明の実施の形態】

以下に、本発明の実施の形態を図面を参照して説明する。

【0046】

（第1の実施の形態）

まず、第1の実施の形態について説明する。第1の実施の形態では、音声で入力された原言語文（以下、表現変換される文のことを原言語文と呼ぶ）を他言語である目的言語文（以下、表現変換された文のことを目的言語文と呼ぶ）に変換する通訳装置について説明する。

【0047】

図1に、本発明の一実施の形態である通訳装置の構成を示す。

【0048】

本実施の形態の通訳装置は、タグ付きコーパス1、依存関係分析手段2、用例DB3、音声認識手段4、重要語抽出手段5、用例文選択手段7、出力文生成手段8、重要語対訳辞書6から構成される。

【0049】

タグ付きコーパス1は、対訳コーパスの各対訳文毎に意図タグを付加したもの

である。

【 0 0 5 0 】

依存関係分析手段 2 は、タグ付きコーパス 1 の各対訳文を重要語の共起関係进行分析することによって用例 D B 3 を作成する手段である。

【 0 0 5 1 】

用例 D B 3 は、原言語文の重要語の共起関係を示す重要語対と目的言語文とを対応つけた用例文を格納したデータベースである。

【 0 0 5 2 】

音声認識手段 4 は、原言語文として入力されてくる音声を音声認識して、単語列候補を出力する手段である。

【 0 0 5 3 】

重要語抽出手段 5 は、音声認識手段 4 から出力されてくる単語列候補を入力し、その単語列候補から予め決められた重要語を抽出する手段である。

【 0 0 5 4 】

用例文選択手段 7 は、入力文中の重要語を組み合わせで作成した重要語対と、用例 D B 3 の各用例文中の重要語対とを比較し、入力文中に最も多く重要語対が含まれている用例文を、最も入力文と文意が類似した用例文であるとみなして選択し、出力する手段である。

【 0 0 5 5 】

出力文生成手段 8 は、出力された用例文の文法的な不自然さを修正し、目的言語文として例えばテキストや音声で出力する手段である。

【 0 0 5 6 】

重要語対訳辞書 6 は、原言語（以下、表現変換される文の言語を原言語と呼ぶ）における重要語とその重要語の目的言語（以下、表現変換された文の言語を目的言語と呼ぶ）における対訳とが対になって格納されている辞書である。

【 0 0 5 7 】

次に、このような本実施の形態の動作を説明する。

【 0 0 5 8 】

本実施の形態の通訳装置は、予め通訳する前に、重要語の内容と重要語群と目

的言語表現パターンとの対応を決定し、重要語の目的言語における対訳を重要語対訳辞書６に記載し、重要語群と目的言語の表現パターンとの対応を用例ＤＢ３に記載しておく必要がある。

【００５９】

そのためには、まず、開発者が、入力文意毎に意図を表す重要語及びそれを用いた表現パターンを手動で決めてもよいし、対訳コーパスの各対訳文毎に意図タグを付け、意図毎に対訳コーパスを分類し、各文意に共通に使用される単語を重要語候補と決め、それらを開発者がチェックすることで半自動的に重要語と表現パターンを決定してもよい。なお、上記の文意とは、同等の意図を表現している一つまたは複数の異なった文を一つにまとめたものをいう。また、対訳コーパスとは、文例のデータベースであり、原言語における文と目的言語における文とが対応つけられた対訳文が多数格納されているものである。

【００６０】

上記のいずれかの方法により決められた重要語と表現パターンとから、重要語対訳辞書、変換用用例ＤＢを作成しておく。図２の（ａ）に、原言語が日本語であり、目的言語が英語である場合、すなわち、本実施の形態の通訳装置が日本語の音声を経英語の音声に通訳する場合に用いられる重要語対訳辞書６の例と用例ＤＢ３の例を示す。

【００６１】

図２の（ａ）の重要語対訳辞書６の例では、日本語の「コーヒー」の対訳として英語の「c o f f e e」が記載されており、日本語の「ミルク」の対訳として英語の「m i l k」が記載されている。また、日本語の「お願い」は、英語では、単一の単語であらわせないで、日本語の「お願い」の対訳とし英語の部分に「＊」が記載されている。

【００６２】

また、図２の（ａ）の用例ＤＢ３の例では、一つの重要語群である「コーヒー」及び「お願い」に目的言語における表現パターンである「I' d l i k e t o c o f f e e p l e a s e . 」が対応していることを示している。また、一つの重要語群である「冷たい」及び「ミルク」及び「あり」に目的言語におけ

る表現パターンである「Do you have a cold milk?」が対応していることを示している。

【0063】

また、図2の(a)の用例DB3の例では、各重要語群で(コーヒー→ミルク)など、重要語が対になって記載されている。このように各重要語は、必ず2ずつ対にして用例DB3に記載されている。このような重要語の対は、重要語の共起関係をあらわしており、依存関係分析手段2により以下に示すようにして作成されたものである。

【0064】

すなわち、まず依存関係分析手段2は、タグ付きコーパス1の原言語文に対して、依存構造分析を行い、各フレーズ毎の依存構造を明らかにする。この依存構造関係で共起関係にある重要語対があれば、その情報を、相当する用例DB3の重要語と表現パターン対に追加する。具体的には、「コーヒーお願い」という対訳コーパスの原言語文に対して、重要語である「コーヒー」と「お願い」とが共起関係にあるので、図2の(a)に示すように、これらを対にした「(コーヒー→お願い)」のように共起関係を付加する。

【0065】

このようにして、タグ付きコーパス1から、図2の(a)に示すような重要語対訳単語辞書6と用例DB3が作成され、また、用例DB3には、各重要語の共起関係が付加される。

【0066】

次に、このようにして、予め作成された用例DB3及び重要語対訳辞書6を用いて通訳を行う際の動作を説明する。

【0067】

通訳の際には、まず、音声認識手段4は、入力されてくる原言語音声の音声認識を行い、認識結果である単語列候補を出力する。そして、この認識結果は、重要語抽出手段5に入力される。例えば、入力されてくる原言語音声が「冷たいコーヒーありますか?」であったとすると、音声認識の結果重要語抽出手段に「冷たいコーヒーありますか?」という認識結果文が出力される。

【0068】

次に、重要語抽出手段5は、予め決められた重要語を認識結果文から抽出し、抽出した重要語を用例文選択手段7に出力する。例えば、「冷たいコーヒーありますか?」という認識結果文に対して、「冷たい」と「コーヒー」と「あり」の3つの重要語を抽出する。

【0069】

そうすると、用例文選択手段7は、重要語抽出手段5から出力されてきた、入力文中の重要語を組み合わせ、入力文中の重要語の対を作成する。そして、入力文中の重要語の対と、用例DB3の各用例文中の重要語対とを比較し、入力文の重要語の対が最も多く含まれている用例文を、最も入力文と文意が類似した用例文であるとみなして選択し、出力する。

【0070】

例えば、入力文中の重要語である「冷たい」と「コーヒー」と「あり」の3つの重要語を組み合わせ、「(冷たい→コーヒー)」、「(コーヒー→あり)」、「(冷たい→あり)」の3通りの入力文中の重要語の対を作成する。

【0071】

そして、用例DB3の一番目の用例文の重要語対と比較する。この用例文の重要語の対は「(コーヒー→お願い)」であるので、入力文中の重要語の対を一つも含んでいない。

【0072】

次に、用例DB3の二番目の用例文の重要語対と比較する。この用例文の重要語の対は、「(冷たい→ミルク)」、「ミルク→あり」のように重要語の対が2つある。そして、入力文中の重要語の対のうち「(冷たい→あり)」は含まないが、「(冷たい→ミルク)」と「(ミルク→あり)」を含んでいる。従って二番目の用例文は、入力文中の重要語の対を2つ含んでいることになる。

【0073】

このようにして入力文中の重要語の対と、用例DB3のすべての用例文の重要語対とを比較した結果、2番目の用例文が、入力文の重要語の対が最も多く含まれている用例文であったとする。そうすると、2番目の用例文を、最も入力文と

文意が類似した用例文であるとみなして選択し、出力する。すなわち、用例文選択手段7から目的言語の表現パターンである「Do you have a cold milk?」が出力される。

【0074】

なお、図1では、出力文生成手段8を図示しているが、出力文生成手段8を備えている必要はなく、上述した用例文選択手段7から出力される目的言語の表現パターンをそのまま外部に出力することにより通訳装置として機能することが出来る。

【0075】

このように本実施の形態では、予め、出力される目的言語文の標準的または簡易的表現パターンとそのパターンに当てはまる重要語群及び重要語間の共起関係とを対応させている用例DB3を作成しておき、入力文から重要語を抽出し、抽出された重要語群と最も類似している重要語群をもつ表現パターンを用例DB3から選択し出力する。従って、もし出力文生成手段8を設けて、出力文生成手段8が変換規則や文生成規則を使用して、代名詞や動詞や助動詞における3人称化や複数化や過去形化などの最適化や、ゼロ代名詞の補間や、文全体の構成の最適化などの文法的な不自然さを修正する場合であっても、その変換規則や文生成規則をコンパクトに構成することが可能となり簡易な処理で高速に表現変換を行うことが出来る。

【0076】

また、重要語以外の部分で誤った文が入力された場合でも、正しく文意が理解できる表現に変換でき、従来のように誤ったまま表現変換結果を出力するという問題を解決することが可能となる。例えば、「冷たいミルクえーつとありますか」や「冷たいのミルクありますか」など、「えーつと」や「の」など重要語以外の部分に誤った部分を含む文が入力された場合であっても、「Do you have a cold milk?」と正しく文意が理解できる表現に変換することが出来る。

【0077】

また、本実施の形態では、用例文選択手段3で用例文が一つだけ選択される場

合について説明したが、抽出された重要語の対と用例ＤＢ３の用例文の重要語の対とを比較した結果、複数の用例文の類似度が同等である場合には、それら複数の用例文を出力すればよい。

【００７８】

なお、予め作成された用例ＤＢ３に記述された表現パターンが、図３に示したように重要語のみで構成されている場合には、出力文生成手段８を設ける必要がないか、または、出力文生成手段８を設けて、変換規則と文生成規則を使用する場合には、さらに変換規則と文生成規則をコンパクトに構成する事が可能となり、効果の高い通訳装置を実現することが可能になる。

【００７９】

（第２の実施の形態）

次に、第２の実施の形態について説明する。

【００８０】

図４に、本発明の一実施の形態である通訳装置の構成を示す。

【００８１】

用例ＤＢ１１は、第１の実施の形態の用例ＤＢとは異なり重要語がクラス化されており、クラスを表す意味コードで置き換えられている点である。

【００８２】

分類語意表１２は、各重要語がどのクラスに属するかを示す意味コードを各重要語に割り当てた一覧表である。

【００８３】

単語クラス化手段１３は、さらに、分類語意表１２の意味コードに従って、用例ＤＢ１１にある重要語、共起関係をなしている重要語対、変換規則に含まれる重要語を意味コードに置き換えて、各重要語をクラス化し、また、分類語意表の意味コードに従って重要語抽出手段５で抽出された重要語が意味コードに置き換えることで、重要語をクラス化する手段である。

【００８４】

用例文選択手段１４は、入力文中の重要語クラスと用例文中の重要語クラス対とを比較し、入力文中に最も多く重要語クラス対が含まれている用例文を、最も

入力文と文意が類似した用例文であるとみなして選択する手段である。

【0085】

出力生成手段15は、選択した用例文の重要語クラスを同じクラスに含まれる入力された重要語に戻し出力する手段である。

【0086】

それ以外は、第1の実施の形態と同様である。

【0087】

次に、このような本実施の形態の動作を説明する。

【0088】

本実施の形態の通訳装置は、第1の実施の形態と同様に、予め通訳する前に、重要語の内容と重要語群と目的言語表現パターンとの対応を決定し、重要語対訳辞書16辞書、用例DB11に記載しておく。そして、依存関係分析手段2は、依存構造関係で共起関係にある重要語対も同様に表現パターン対に追加しておく。

【0089】

さらに、単語クラス化手段13は、分類語意表12の意味コードに従って、用例DB11にある重要語、共起関係をなしている重要語対、変換規則に含まれる重要語を意味コードに置き換えて、各重要語をクラス化しておく。

【0090】

図5の(a)に、分類語意表12の例を示し、図5の(b)に用例DB11の例を示す。

【0091】

図5の(a)に示す分類語意表12では、「コーヒー」などの飲み物を示す単語の意味コードが100であり、「冷たい」などの飲み物の状態を表す意味コードが200

である。このように分類語意表12の各重要語は、それぞれ意味コードを付与されることによってクラス化されている。

【0092】

また、図5の(b)に示す用例DB11の例では、分類語意表12で各重要語に付与された各々の意味コードを用いて、第1の実施の形態で説明した図2の用

例 D B 3 の例の中に現れる重要語が意味コードで表されている。

【 0 0 9 3 】

次に、このようにして、予め作成された用例 D B 1 1 及び重要語対訳辞書 1 6 を用いて通訳を行う際の動作を説明する。

【 0 0 9 4 】

通訳の際には、まず、音声認識手段 4 に原言語音声が入力され、音声認識手段 4 は、入力されてくる原言語音声を音声認識し、認識結果である単語列候補を出力する。そして、この認識結果は、重要語抽出手段 5 に入力される。

【 0 0 9 5 】

次に、重要語抽出手段 5 は、予め決められた重要語を認識結果文から抽出する。そして、単語クラス化手段 1 3 は、分類語意表 1 2 の意味コードに従って抽出された重要語を意味コードに置き換えることで、重要語を単語クラス化する。

【 0 0 9 6 】

次に、用例文選択手段 1 4 は、単語クラス化手段 1 3 で単語クラス化された入力文中の重要語のクラスを組み合わせ、入力文中の重要語のクラスの対を作成する。そして、入力文中の重要語のクラスの対と、用例 D B 1 1 中の各用例文中の重要語のクラス対とを比較し、入力文中に最も多く重要語のクラス対が含まれている用例文を、最も入力文と文意が類似した用例文であるとみなして選択する。

【 0 0 9 7 】

出力文生成手段 1 5 は、用例文生成手段 1 4 が選択した用例文の重要語のクラスを同じクラスに含まれる入力された重要語に戻し、重要語対訳辞書 1 6 を用いて入力された重要語の対訳に置き換えて出力する。

【 0 0 9 8 】

また、本実施の形態では、用例文選択手段 2 1 で用例文が一つだけ選択される場合について説明したが、抽出された重要語のクラスの対と用例 D B 3 の用例文のクラスの対とを比較した結果、複数の用例文の類似度が同等である場合には、複数の用例文を出力すればよい。

【 0 0 9 9 】

このように本実施の形態によれば、予め、出力される目的言語文の標準的または簡易的表現パターンとそのパターンに当てはまる重要語クラス群及び重要語クラス群間の共起関係とを対応させている用例ＤＢ１１を作成しておき、入力文から重要語を抽出し、抽出された重要語を分類語意表１２に従ってクラス化し、クラス化された重要語群と最も類似している重要語クラス群をもつ表現パターンを用例文ＤＢ１１から選択し、クラス化された重要語をもとの重要語に戻して出力することで、変換規則や文生成規則を第１の実施の形態と同様にコンパクトに構成することが可能となり簡易な処理で表現変換を行うことが可能である。

【０１００】

また、重要語以外の部分で誤った文が入力された場合でも、正しく文意が理解できる表現に変換でき、従来のように誤ったまま表現変換結果を出力するという問題を解決することが可能となる。

【０１０１】

さらに、単語クラス化することで、用例文にはない新たな入力文に対しても正しい翻訳結果を出力することが可能となるため、少ない用例文データベースで多くの入力文に対応できる通訳装置を実現することが出来る。

【０１０２】

なお、本実施の形態では、予め通訳する前に、重要語の内容と重要語群と目的言語表現パターンとの対応を決定し、重要語対訳辞書１６辞書、用例ＤＢ１１に記載し、依存構造関係で共起関係にある重要語対も同様に表現パターン対に追加した後、単語クラス化手段１３は、分類語意表１２の意味コードに従って、用例ＤＢ１１にある重要語、共起関係をなしている重要語対、変換規則に含まれる重要語を意味コードに置き換えて、各重要語をクラス化しておくとして説明したが、これに限らない。まず、予め通訳する前に分類語意表１２を利用して、タグ付きコーパス１の各文について、その文から重要語を意味コードに置き換え、それらの意味コードのうちの、所定の数の組み合わせの中から共起関係のある意味コードの組み合わせを特定して、その特定した意味コードの組み合わせとその意味コードを選択する対象となった文の変換後の表現とを予め対応させておいても構わない。

【 0 1 0 3 】

(第 3 の実施の形態)

次に、第 3 の実施の形態について説明する。第 3 の実施の形態でも第 1 及び第 2 の実施の形態と同様に通訳装置を例にして説明する。

【 0 1 0 4 】

図 6 に、本実施の形態の通訳装置の構成を示す。

【 0 1 0 5 】

用例文選択手段 2 1 は、第 1 の実施の形態の用例文選択手段 7 の機能に加え、さらに、用例 D B 3 にある重要語対と抽出された重要語群とを比較し、抽出された重要語の中で、他の重要語のどれとも共起関係による重要語対にならない重要語を誤認識単語と推定して用例文を選択する手段である。

【 0 1 0 6 】

それ以外は、第 1 の実施の形態と同様である。

【 0 1 0 7 】

次に、このような本実施の形態の動作を説明する。

【 0 1 0 8 】

本実施の形態の通訳装置は、第 1 の実施の形態と同様にして、予め通訳する前に、重要語の内容と重要語群と目的言語表現パターンとの対応を決定し、重要語対訳辞書 6、用例 D B 3 に記載し、重要語の共起関係も同様に用例 D B 3 に付加する。

【 0 1 0 9 】

次に、通訳の際の動作を説明する。

【 0 1 1 0 】

通訳の際には、まず、音声認識手段 4 に原言語音声が入力され、音声認識手段 4 は、原言語音声の音声認識を行い、認識結果である単語列候補が出力される。

【 0 1 1 1 】

重要語抽出手段 5 は、音声認識手段 4 から出力された認識結果を入力し、予め決められた重要語を認識結果文から抽出する。

【 0 1 1 2 】

次に、用例文選択手段21で、用例DB3にある重要語対と抽出された重要語群とを比較し、抽出された重要語の中で、他の重要語のどれとも共起関係による重要語対にならない重要語を誤認識単語と推定する。

【0113】

そして、用例文選択手段21は、誤認識であるとみなされた単語を除いた残りの重要語を用いて、第1の実施の形態と同様に最も入力文と意図が類似した用例文を選択し出力する。

【0114】

図7に例を示す。音声認識手段4には「熱いミルク」と入力されたにもかかわらず、音声認識手段4が「青いミルク」と誤認識をした例である。すなわち、音声認識手段4には、「熱いミルクはありますか」という音声が入力されたが、音声認識手段4が誤認識したために、「青いミルクはありますか」という認識結果文を出力した場合である。

【0115】

また、図7の用例DB3の例では、「コーヒー」、「お願い」という一つの重要語群が「(コーヒー→お願い)」という重要語の対になっており、目的言語における表現パターンである「C o f f e e p l e a s e . 」が対応していることを示している。また、一つの重要語群である「熱い」及び「ミルク」及び「あり」が「熱い→ミルク」、及び「ミルク→あり」という重要語の対を持ち、目的言語における表現パターンである「D o y o u h a v a a h o t m i l k ? 」が対応していることを示している。

【0116】

上記のように音声認識手段4が誤認識した場合、重要語抽出手段5は、「青いミルクありますか」という認識結果文を入力し、重要語として、「青い」、「ミルク」、「あり」の3つの重要語を抽出する。

【0117】

このような場合、用例文選択手段21は、重要語抽出手段5が抽出した「青い」、「ミルク」、「あり」の3つの重要語を組み合わせ「(青い→ミルク)」、「(青い→あり)」、「(ミルク→あり)」の3つの重要語の対を作成し、こ

れらを用例DB3の重要語群と比較する。

【0118】

すなわち、まず、用例DB3に記述されている一番目の用例文の重要語の対である「(コーヒー→お願い)」と、組み合わせた3つの重要語の対とを比較する。その結果一番目の用例文の重要語の対である「(コーヒー→お願い)」と一致する重要語対は存在しない。

【0119】

次に、用例DB3に記述されている二番目の用例文の重要語の対である「(熱い→ミルク)」、及び「(ミルク→あり)」と、組み合わせた3つの重要語の対とを比較する。その結果、組み合わせた重要語の対のうち「(青い→ミルク)」及び「(青い→あり)」は、二番目の用例文の重要語の対のいずれとも一致しないが、「(ミルク→あり)」は、二番目の用例文の重要語の対と一致している。

【0120】

すなわち、用例DB3の一番目の用例文の重要語の対よりも二番目の用例文の重要語の対の方が、組み合わせた重要語の対と類似度が高い。従って二番目の用例文を選択する。

【0121】

さらに、二番目の用例文の重要語の対で組み合わせた重要語の対と不一致となった方の、重要語の対に属する重要語で、組み合わせた重要語の対と一致している方の、重要語の対に含まれていない重要語については、二番目の用例文からその重要語の変換後の部分を取り除いてから二番目の用例文を出力する。すなわち、「熱い」がそのような重要語であるので、重要語対訳辞書6を参照することにより、「Any hot milk?」から「熱い」に対応する変換語の表現である「hot」を取り除いて「Any milk?」という文を出力する。

【0122】

このように、「(青い→ミルク)」、「(青い→あり)」、「(ミルク→あり)」の3つの重要語の対と、用例DB3の共起関係すなわち重要語の対とを比べた場合、「熱い」と共起関係を持つ単語が認識結果の中に存在しないため、「熱い」が誤認識されたと推定し、上述したように「Any milk?」という文

を出力する。

【0123】

また、本実施の形態では、用例文選択手段21で用例文が一つだけ選択される場合について説明したが、抽出された重要語の対と用例DB3の用例文の重要語の対とを比較した結果、複数の用例文の類似度が同等である場合には、それら複数の用例文を出力すればよい。

【0124】

なお、本実施の形態における用例文選択手段21の機能は、上記のものに限らず、用例文選択手段21の機能を誤認識選択手段と用例文選択手段との2段階の機能に分けても構わない。この場合、誤認識単語推定手段は、認識結果から抽出された重要語群と用例DB3に記載されている重要語の対とを比較することによって、誤認識の単語を推定する手段であり、用例文選択手段は、第1の実施の形態の用例文選択手段7と同等の機能を持つ手段である。

【0125】

この場合、通訳の際には、誤認識単語推定手段21は、作成した重要語の対を用例DB3の各用例文の重要語の対と比較する。そして、抽出された重要語の中で他の重要語のどれとも共起関係による重要語にならない重要語を誤認識単語と推定する。次に、用例文選択手段で、誤認識単語推定手段で誤認識であるとみなされた単語を除いた残りの重要語を用いて、第1の実施の形態と同様に最も入力文と意図が類似した用例文を選択し、出力する。このようにすれば、上述したように音声認識手段4での認識結果に誤認識が含まれており、音声認識手段4で「青いミルクありますか」という認識結果が出力された場合でも、誤認識単語推定手段と用例文選択手段で処理されることにより、上述した場合と同様に「Any milk?」を出力することが出来る。

【0126】

なお、上記第1～第3の各実施の形態では、原言語文が日本語の文であり、目的言語文が英語の文の場合、すなわち、通訳装置が日本語の文を英語の文に通訳する場合について説明したが、原言語文や目的言語文はこれ以外の言語の文であっても構わない。

【0127】

以下に、通訳装置が英語の文を日本語の文に翻訳する場合と、通訳装置が中国語の文を日本語の文に翻訳する場合について、上記実施の形態との相違点を中心に説明する。

【0128】

まず、通訳装置が英語の文を日本語の文に翻訳する場合、上記実施の形態で説明した図2、図5、及び図7は、それぞれ図10、図11、及び図12に置き換えればよい。

【0129】

すなわち、図10の(a)はこの場合の重要語対訳単語辞書6と用例DB3の例であり、図10の(b)は、この場合のタグ付きコーパス1の例である。

【0130】

また、図11の(a)は、この場合の分類語意表12の例であり、図11の(b)は、この場合の用例DB11の例である。

【0131】

また、図12は、この場合の用例DB3の例である。

【0132】

このように、図2、図5、及び図7を、それぞれ図10、図11、及び図12に置き換えることによって、上記各実施の形態が英語の文を日本語の文に通訳する場合にも適用できることは明らかである。

【0133】

また、通訳装置が中国語の文を日本語の文に翻訳する場合、上記実施の形態で説明した図2、図5、及び図7は、それぞれ図13、図14、及び図15に置き換えればよい。

【0134】

すなわち、図13の(a)はこの場合の重要語対訳単語辞書6と用例DB3の例であり、図13の(b)は、この場合のタグ付きコーパス1の例である。

【0135】

また、図14の(a)は、この場合の分類語意表12の例であり、図14の(b)

b) は、この場合の用例 D B 1 1 の例である。

【 0 1 3 6 】

また、図 1 5 は、この場合の用例 D B 3 の例である。

【 0 1 3 7 】

このように、図 2、図 5、及び図 7 を、それぞれ図 1 3、図 1 4、及び図 1 5 に置き換えることによって、上記各実施の形態が中国語の文を日本語の文に通訳する場合にも適用できることは明らかである。

【 0 1 3 8 】

さらに、本実施の形態では、依存関係分析手段 2 は、共起関係にある重要語を重要語の対とすると説明したが、これに限らず、依存関係分析手段 2 は、依存関係にある重要語を重要語の対としても構わない。

【 0 1 3 9 】

なお、本実施の形態のタグ付きコーパス 1 は本発明のコーパスの例であり、本実施の形態の重要語の対は本発明の重要語の組み合わせの例であり、本実施の形態の依存関係分析手段 2 は本発明の対応付け手段の例であり、本実施の形態の重要語抽出手段 5、用例文選択手段 7 は本発明の変換手段の例であり、本実施の形態の分類語意表 1 2 は本発明のクラス化情報の例であり、本実施の形態の依存関係分析手段 2、単語クラス化手段 1 3 は本発明の対応付け手段の例であり、本実施の形態の重要語抽出手段 5、用例文選択手段 1 4 は本発明の変換手段の例であり、本実施の形態の重要語抽出手段 5、用例文選択手段 2 1 は本発明の変換手段の例であり、本実施の形態の意味コードは本発明のクラス名の例である。

【 0 1 4 0 】

さらに、本発明の所定の数とは、本実施の形態における 2 に限らない。所定の数が 1 である場合でも構わない。この場合、重要語を組み合わせで重要語の対と比較する代わりに、重要語単独で比較するものとする。また、所定の数が 3 である場合でも構わない。この場合、3 つの重要語の組み合わせで比較するものとする。また、所定の数が 2 及び 3 の場合であっても構わない。この場合、重要語の対と 3 つの重要語の組み合わせの両方を用いて比較するものとする。要するに本発明の所定の数とは、任意の正の整数または複数の異なった任意の正の整数であ

りさえすればよい。

【 0 1 4 1 】

さらに、本発明の表現変換装置は、本実施の形態における通訳装置に限らず、日本語のテキストを入力して英語のテキストに変換して出力するなどの異なった言語にテキストを変換して出力する翻訳装置、書き言葉を話し言葉に変換する文型変換装置、複雑なまたは冗長な文章を要約して出力する要約文作成装置など、要するに本発明の表現変換装置は、入力文を異なった表現に変換する装置でありさえすればよい。

【 0 1 4 2 】

さらに、本発明の表現変換方法は、本実施の形態における通訳を行う方法に限らず、日本語のテキストを入力して英語のテキストに変換して出力するなどの異なった言語にテキストを変換して出力する翻訳、書き言葉を話し言葉に変換する文型変換、複雑なまたは冗長な文章を要約して出力する要約文作成など、要するに本発明の表現変換方法は、入力文を異なった表現に変換する方法でありさえすればよい。

【 0 1 4 3 】

このように予め、出力される目的言語文の標準的または簡易的表現パターンとそのパターンに当てはまる重要語群及び重要語群間の共起関係とを対応させている用例文データベースを作成しておき、入力文から重要語し、抽出された重要語群と最も類似している重要語群をもつ表現パターンを用例文データベースから選択し出力することで、変換規則や文生成規則をコンパクトに構成することが可能となり簡易な処理で表現変換を行うことを可能にすることが出来る。

【 0 1 4 4 】

また、重要語以外の部分で誤った文が入力された場合でも、正しく文意が理解できる表現に変換でき、従来のように誤ったまま表現変換結果を出力するという問題を解決することが可能となる。

【 0 1 4 5 】

さらに、たとえ重要語が誤っていても、その誤りの程度に応じて、正しく文意が理解できる表現に変換するか、意味が理解できなかったことをユーザに告知す

ることが可能となり、従来のように誤ったまま表現変換結果を出力するという問題を解決することが可能となる。

【０１４６】

さらに、本発明は、上述した本発明の表現変換装置の全部または一部の手段（または、装置、素子、回路、部等）の機能をコンピュータにより実行させるためのプログラムであって、コンピュータと協働して動作するプログラムである。

【０１４７】

さらに、本発明は、上述した本発明の表現変換方法の全部または一部のステップ（または、工程、動作、作用等）の動作をコンピュータにより実行させるためのプログラムであって、コンピュータと協働して動作するプログラムである。

【０１４８】

なお、本発明の一部の手段（または、装置、素子、回路、部等）、本発明の一部のステップ（または、工程、動作、作用等）には、それらの複数の手段またはステップの内の、幾つかの手段またはステップを意味し、あるいは、一つの手段またはステップの内の、一部の機能または一部の動作を意味するものである。

【０１４９】

また、本発明のプログラムを記録した、コンピュータに読みとり可能な記録媒体も本発明に含まれる。

【０１５０】

また、本発明のプログラムの一利用形態は、コンピュータにより読み取り可能な記録媒体に記録され、コンピュータと協働して動作する態様であっても良い。

【０１５１】

また、本発明のプログラムの一利用形態は、伝送媒体中を伝送し、コンピュータにより読みとられ、コンピュータと協働して動作する態様であっても良い。

【０１５２】

また、記録媒体としては、ROM等が含まれ、伝送媒体としては、インターネット等の伝送媒体、光・電波・音波等が含まれる。

【０１５３】

また、上述した本発明のコンピュータは、CPU等の純然たるハードウェアに

限らず、ファームウェアや、ＯＳ、更に周辺機器を含むものであっても良い。

【０１５４】

なお、以上説明した様に、本発明の構成は、ソフトウェア的に実現しても良いし、ハードウェア的に実現しても良い。

【０１５５】

以上詳述したように、本実施の形態によれば、入力文から重要語を抽出し、抽出された重要語を用いて同じ文意を表す標準的または簡易的表現文に変換し出力することで、変換規則や文生成規則をコンパクトに構成することが可能となり簡易な処理で表現変換を行うことを可能にする。また、重要語以外の部分で誤った文が入力された場合でも、正しく文意が理解できる表現に変換でき、従来のように誤ったまま表現変換結果を出力するという問題を解決することが可能となる。

【０１５６】

また、入力文に含まれる一部の内容語の単語もしくはその変換された単語を重要語として抽出し、この重要語と入力文から推定される文意により決まる表現との組み合わせからなる標準的または簡易的表現文を生成することで、変換規則や文生成規則をコンパクトに構成することが可能となり簡易な処理で表現変換を行うことを可能にする。

【０１５７】

また、重要語以外の部分で誤った文が入力された場合でも、正しく文意が理解できる表現に変換でき、従来のように誤ったまま表現変換結果を出力するという問題を解決することが可能となる。

【０１５８】

また、入力文から重要語を抽出し、抽出された重要語の共起依存関係または共起関係から文意を推定し、重要語またはその対訳単語と推定された意図から予め決められている単語のみの組み合わせから標準的または簡易的表現を生成することで、変換規則や文生成規則をコンパクトに構成することが可能となり簡易な処理で表現変換を行うことを可能にする。

【０１５９】

また、重要語以外の部分で誤った文が入力された場合でも、正しく文意が理解

できる表現に変換でき、従来のように誤ったまま表現変換結果を出力するという問題を解決することが可能となる。

【 0 1 6 0 】

また、予め、出力される標準的または簡易的表現パターンとそのパターンに当てはまる重要語群及び重要語群間の依存関係または共起関係とを対応させている用例DBを作成しておき、入力文から重要語群を抽出し、抽出された重要語群と最も類似している重要語群をもつ表現パターンを用例DBから選択し、選択された用例文を出力することにより、上記効果に加え、さらに現実に入力される文の種類やドメインや文型に忠実に確度よく表現変換を行うことを可能にする。

【 0 1 6 1 】

また、用例DBに記載されている表現パターンが重要語または重要語の対訳単語のみで構成されていることにより、さらに強化することが可能である。

【 0 1 6 2 】

また、用例DBに記載されている表現パターン、重要語群及び重要語群間の依存関係または共起関係の重要語の記述を重要語を含む単語クラスの記述とすることで、上記効果に加え、用例文DBに含まれていない重要語が入力されても妥当な用例文を選択可能となり、より多種多様な入力文に対応できる表現変換が可能となる。

【 0 1 6 3 】

また入力文から重要語群を抽出し、抽出された重要語間の関係から入力誤りである単語を推定し、推定された単語を除いた残りの重要語から文意を推定し、推定される文意により決まる単語の組み合わせから標準的または簡易的表現を生成することで、上記効果に加えて、たとえ重要語が誤っていても、その誤りの重度に応じて、正しく文意が理解できる表現に変換するか、意味が理解できなかったことをユーザに告知することが可能となり、従来のように誤ったまま表現変換結果を出力するという問題を解決することが可能となる。

【 0 1 6 4 】

【発明の効果】

本発明は、コンパクトな構成であり、かつ高速に処理可能な表現変換方法、

表現変換装置、及びプログラムを提供することが出来る。

【０１６５】

また、本発明は、音声認識の誤りなどにより入力文中の重要語以外で認識誤りを起こしても、出力文の質に悪影響を及ぼすことなく意図が正しく伝わる結果を出力できる表現変換方法、表現変換装置、及びプログラムを提供することが出来る。

【０１６６】

また、本発明は、音声認識の誤りなどにより入力文中で認識誤りを起こしても、従来のように全く文意が伝わらない結果を出力することを回避することが出来る表現変換方法、表現変換装置、及びプログラムを提供することが出来る。

【図面の簡単な説明】

【図１】

本発明の第１の実施の形態の通訳装置を示す図

【図２】

(a) 本発明の第１の実施の形態で使用する重要語対訳辞書と用例ＤＢとの例を示す図

(b) 本発明の第１の実施の形態で使用するタグ付きコーパスの例を示す図

【図３】

本発明の第１の実施の形態で使用する用例ＤＢを示す図

【図４】

本発明の第２の実施の形態の通訳装置を示す図

【図５】

(a) 本発明の第２の実施の形態で使用する分類語意表を示す図

(b) 本発明の第２の実施の形態で使用する用例ＤＢを示す図

【図６】

本発明の第３の実施の形態の通訳装置を示す図

【図７】

本発明の第３の実施の形態で使用する用例ＤＢを示す図

【図８】

従来の通訳装置の構成を示す図

【図 9】

- (a) 従来の対訳発声文例の一例を示す図
- (b) 従来の対訳フレーズの一例を示す図
- (c) 従来の対訳フレーズ辞書の一例を示す図
- (d) 従来のフレーズ間規則の一例を示す図
- (e) 従来の分類語意表の一例を示す図

【図 1 0】

- (a) 本発明の第 1 の実施の形態で英語の文を日本語の文に通訳する場合に使用する重要語対訳辞書と用例 D B との例を示す図
- (b) 本発明の第 1 の実施の形態で英語の文を日本語の文に通訳する場合に使用するタグ付きコーパスの例を示す図

【図 1 1】

- (a) 本発明の第 2 の実施の形態で英語の文を日本語の文に通訳する場合に使用する分類語意表を示す図
- (b) 本発明の第 2 の実施の形態で英語の文を日本語の文に通訳する場合に使用する用例 D B を示す図

【図 1 2】

本発明の第 3 の実施の形態で英語の文を日本語の文に通訳する場合に使用する用例 D B を示す図

【図 1 3】

- (a) 本発明の第 1 の実施の形態で中国語の文を日本語の文に通訳する場合に使用する重要語対訳辞書と用例 D B との例を示す図
- (b) 本発明の第 1 の実施の形態で中国語の文を日本語の文に通訳する場合に使用するタグ付きコーパスの例を示す図

【図 1 4】

- (a) 本発明の第 2 の実施の形態で中国語の文を日本語の文に通訳する場合に使用する分類語意表を示す図
- (b) 本発明の第 2 の実施の形態で中国語の文を日本語の文に通訳する場合に

使用する用例 D B を示す図

【図 1 5】

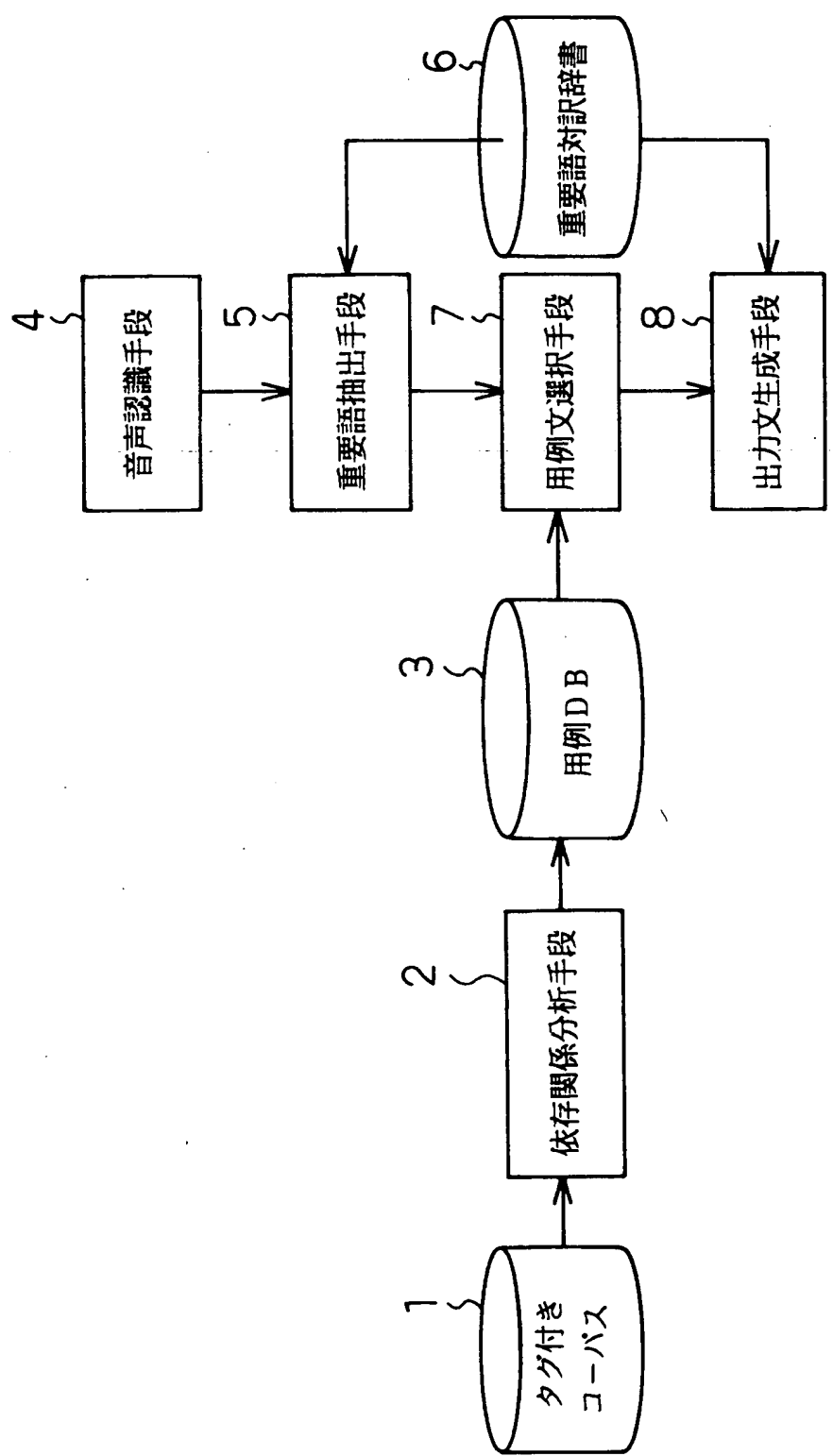
本発明の第 3 の実施の形態で中国語の文を日本語の文に通訳する場合に使用する用例 D B を示す図

【符号の説明】

- 1 タグ付きコーパス
- 2 依存関係分析手段
- 3, 1 1 用例データベース
- 4 音声認識手段
- 5 重要語抽出手段
- 6, 1 6 重要語対訳辞書
- 7, 1 4 用例文選択手段
- 8, 1 5 出力文生成手段
- 1 2 分類語意表
- 1 3 単語クラス化手段
- 2 1 誤認識単語推定手段

【書類名】 図面

【図1】



【図2】

(a) 重要語対訳辞書例

コーヒー	: coffee
お願い	: *
ミルク	: milk
冷たい	: cold
あり	: *

用例DB例

依存関係	: 表現ボタン
(コーヒー→お願い)	: I'd like to coffee please.
(冷たい→ミルク) (ミルク→あり)	: Do you have a cold milk ?
...	

(b)

タグ付きコーパス例

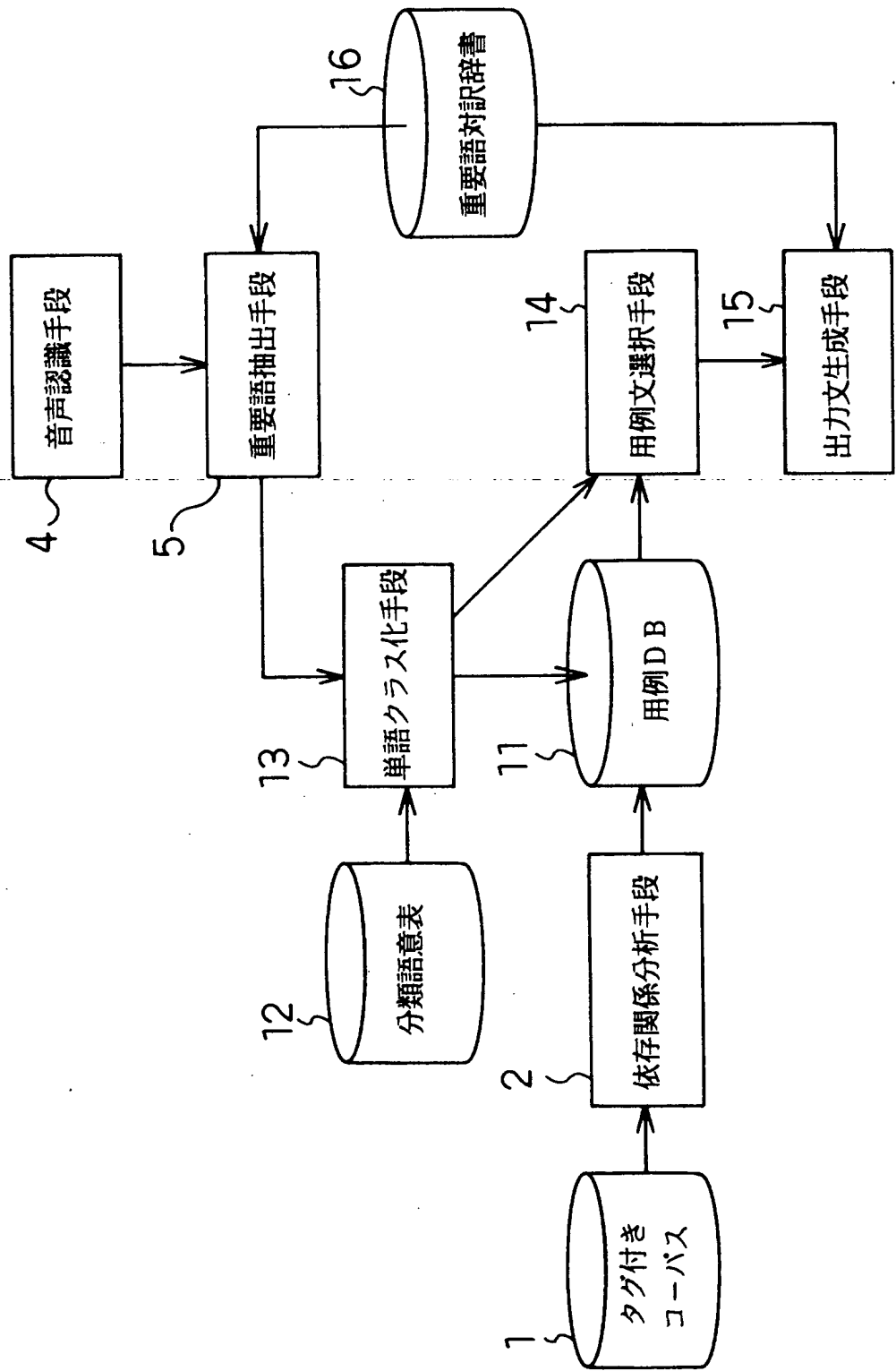
コーヒー (一般名詞) を (格助詞) お (接頭詞) 願い (サ変名詞) し (動詞) ます (助動詞)
冷たい (形容詞) ミルク (普通名詞) は (係助詞) あり (動詞) ます (助動詞) か (終助詞)
...

【図 3】

用例 D B 例

依存関係	: 表現ボタン
(コーヒー→お願い)	: Coffee please.
(冷たい→ミルク) (ミルク→あり)	: Any cold milk ?
...	

【図4】



【図5】

(a) 分類語意表

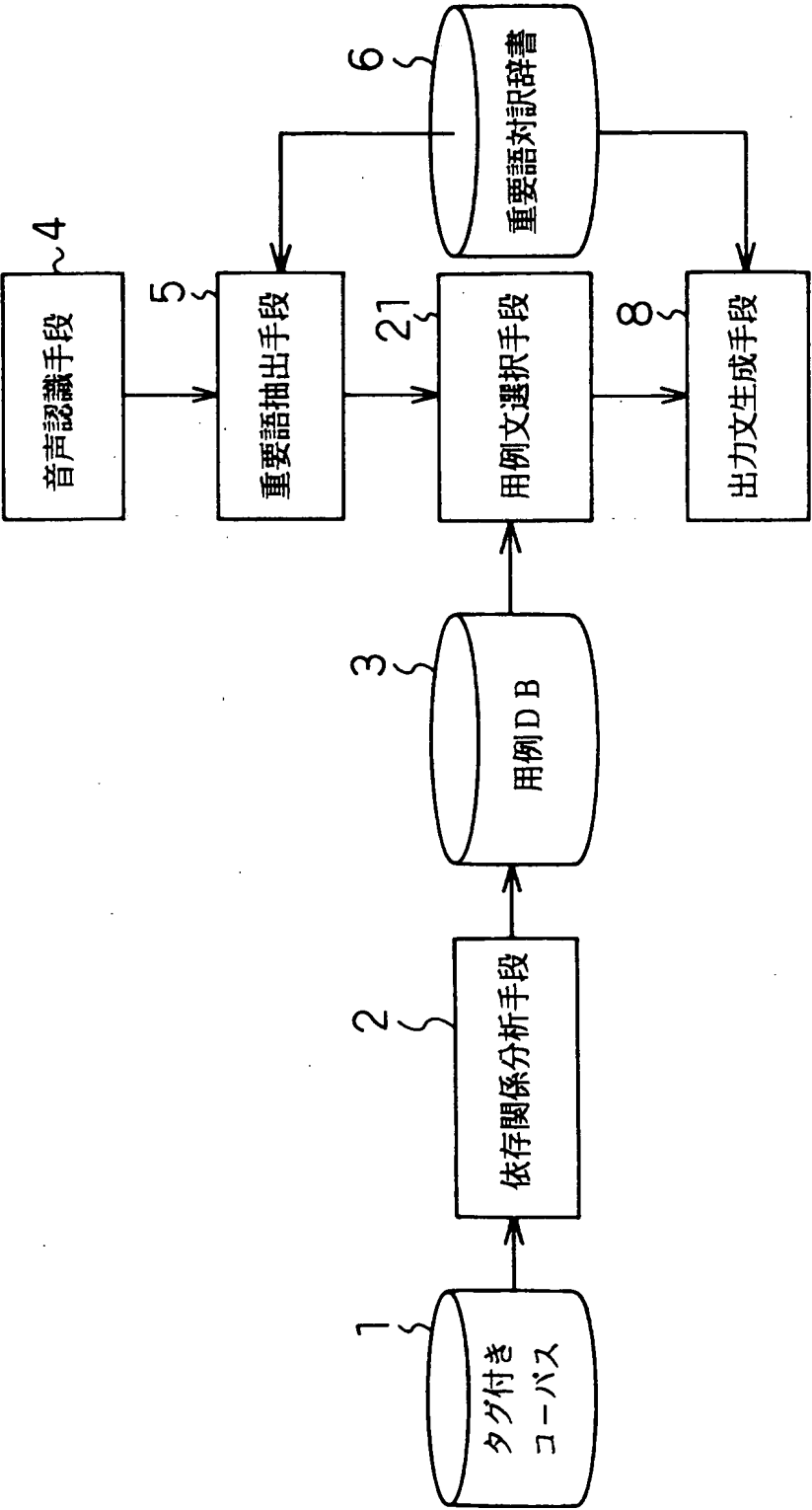
コーヒー (一般名詞)	100
ミルク (一般名詞)	100
紅茶 (一般名詞)	100
...	
冷たい (形容詞)	200
熱い (形容詞)	200
...	

(b)

用例DB例

重要語	:	依存関係	:	用例文
①100 ②お願い	:	(①→②)	:	I'd like to ② please.
①200 ②100 ③あり	:	(①→②) (②→③)	:	Do you have a ① ② ?
...				

【図6】



【図7】

入力文 : 熱いミルクはありますか
認識結果文 : 青いミルクはありますか

用例DB例

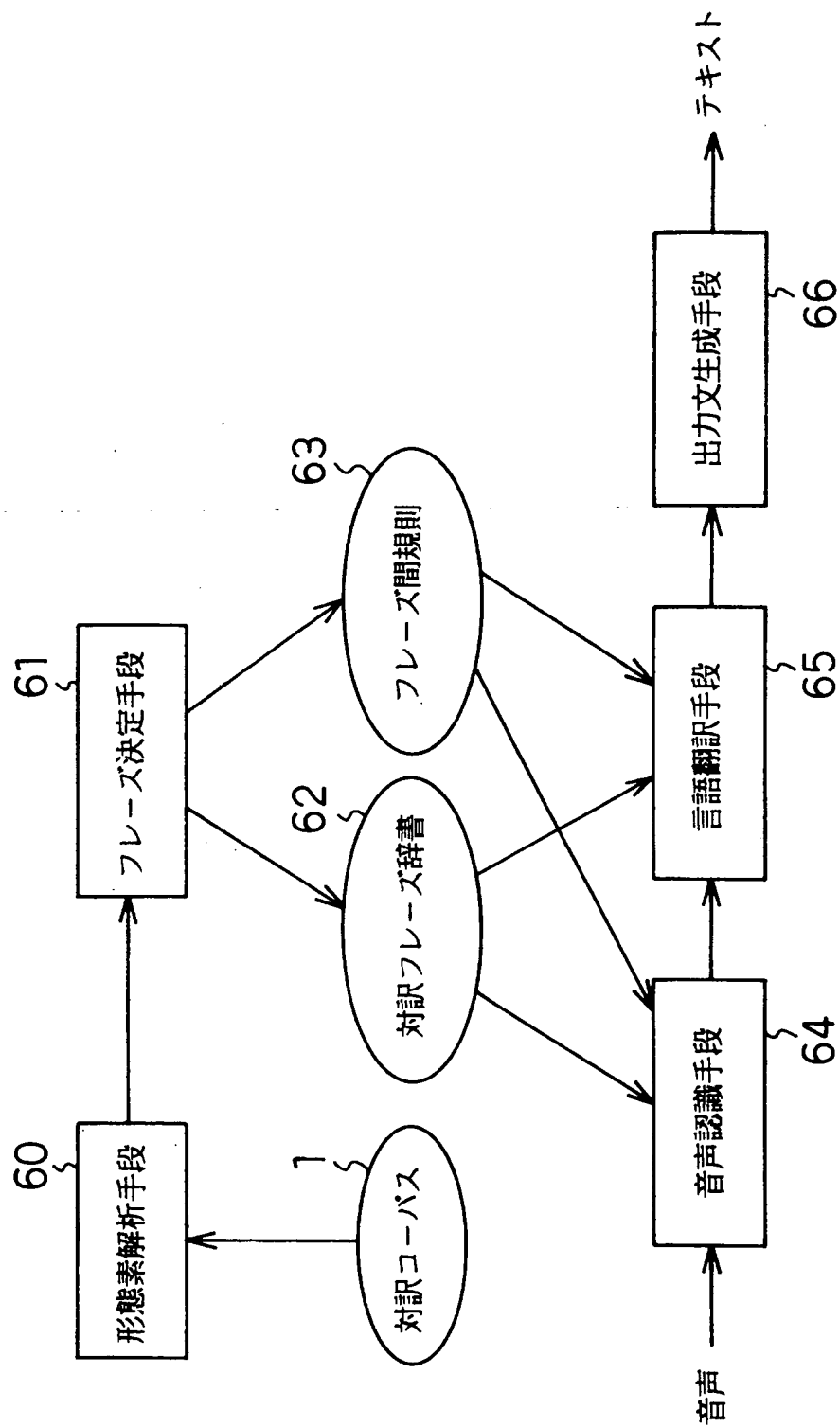
依存関係	: 用例文
(コーヒー→お願い)	: Coffee please.
(熱い→ミルク) (ミルク→あり)	: Any hot milk ?
...	

抽出された重要語: 青い、ミルク、あり

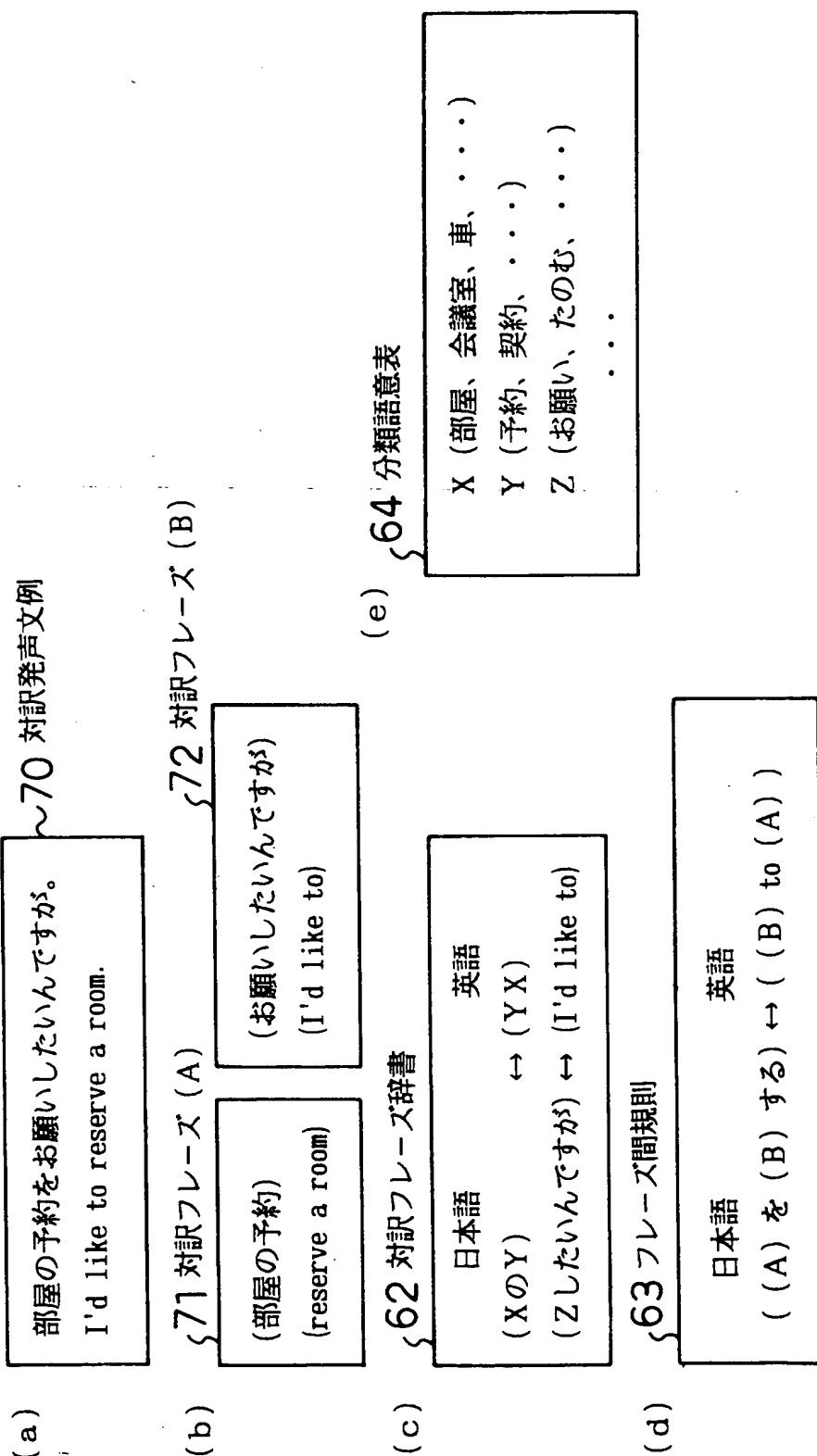
上記重要語と用例DBとの依存関係比較結果:

- (青い、ミルク) ×
- (青い、あり) ×
- (ミルク、あり) ○

【図8】



【図9】



【図10】

(a) 重要語対訳辞書例

coffee	:	コーヒー
please	:	お願い
milk	:	ミルク
cold	:	冷たい
have	:	あり

用例D B例

依存関係	: 表現ボタン
(coffee→please)	: コーヒーをお願いします。
(cold→milk) (milk→have)	: 冷たいミルクはありますか。
...	

(b)

タグ付きコーパス例

I (pronoun) 'd (auxiliary verb) like (verb) to (determiner) coffee (common noun) please (adverb) ...
--

【図11】

(a)

分類語意表

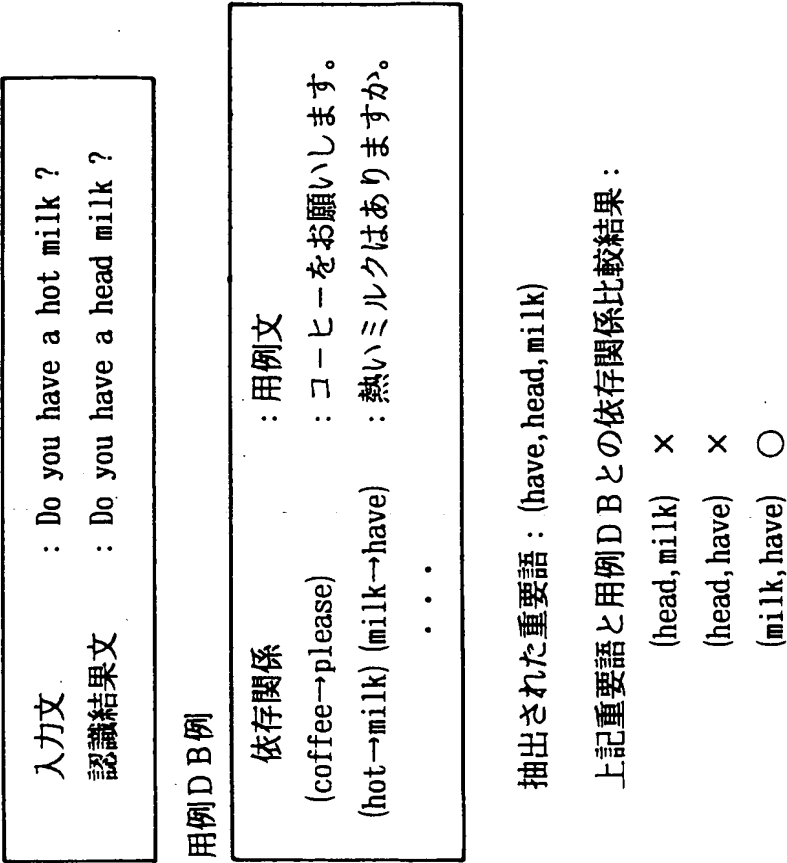
coffee (common noun)	100
milk (common noun)	100
tea (common noun)	100
...	
cold (adjective)	200
hot (adjective)	200
...	

(b)

用例DB例

重要語	依存関係	用例文
①100 ② please	: (①→②)	: ①を②します。
①200 ②100 ③ have	: (①→②) (②→③)	: ①②は③ますか。
...		

【図12】



【図13】

(a)

重要語対訳辞書例

咖啡 : コーヒー
 要 : 下さい
 日文 : 日本語
 菜单 : メニュー
 有 : あり

用例D B例

依存関係 : 表現ボタン
 (咖啡→要) : コーヒーを下さい。
 (日文→菜单)(菜单→有) : 日本語のメニューはありますか。
 ...

(b)

タグ付きコーパス例

要 (動) | 咖啡 (名 詞) |
 ...

【図14】

(a) 分類語意表

咖啡	(名)	100
牛奶	(名)	100
红茶	(名)	100
冷	(形)	200
热	(形)	200
日文	(名)	300
菜单	(名)	400
...		

(b)

用例D B例

重要語		依存関係		用例文
①100 ②要		: (①→②)		: ①を②
①300 ②400 ③有		: (①→②) (②→③)		: ①の②は③ですか。
...				

【図15】

入力文 : 有日文菜单吗
認識結果文 : 有日没菜单吗

用例D B例

依存関係	: 用例文
(咖啡→要)	: コーヒーを下さい。
(日文→菜单)(菜单→有)	: 日本語のメニューはありますか。
...	

抽出された重要語: (有, 日没, 菜单)

上記重要語と用例D Bとの依存関係比較結果:

(日没, 有) ×
(日没, 菜单) ×
(菜单, 有) ○

【書類名】 要約書

【要約】

【課題】 コンパクトな構成であり、かつ高速に処理可能であり、認識誤りを起こしても、意図が正しく伝わる結果を出力できる表現変換装置を提供すること。

【解決手段】 コーパス1の各文について、その文から重要語を選んで、それらの重要語のうちの、所定の数の組み合わせの中から共起関係のある重要語の組み合わせを特定して、その特定した重要語の組み合わせとその重要語を選択する対象となった文の変換後の表現とを予め対応させておく依存関係分析手段2と、入力されてきた文から予め決まっている重要語を選んで、組み合わせ、その組み合わせた重要語の組み合わせと、前記各文の予め前記特定した重要語の組み合わせとを対比し、その対比結果が一致もしくは類似度の高い文を1または2以上選んで、その文の変換後の表現を出力する重要語抽出手段5及び用例文選択手段7とを備える。

【選択図】 図1